

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2000-148736

(43)Date of publication of application : 30.05.2000

(51)Int.Cl. G06F 17/21

(21)Application number : 10-321630

(71)Applicant : INTERNATL BUSINESS MACH CORP <IBM>

(22)Date of filing : 12.11.1998

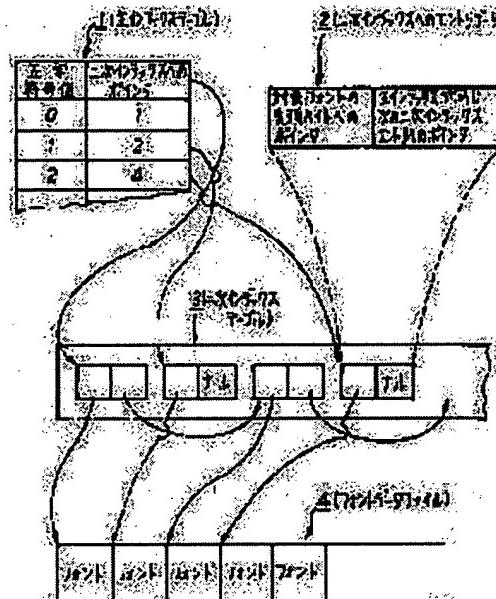
(72)Inventor : KIDO AKIO

(54) METHODS FOR FONT ACQUISITION, REGISTRATION, DISPLAY, AND PRINTING, METHOD FOR HANDLING DOCUMENT HAVING VARIANT FONTS, AND RECORDING MEDIUM THEREOF

(57)Abstract:

PROBLEM TO BE SOLVED: To enable a creator of electronic documents to use a large variety of characters beyond the restriction of the number of characters of conventional character codes by shifting the position of entry data to a secondary index by a variant type number and obtaining a font image as an objective variant-type font.

SOLUTION: On the basis of the variant type number and the character code value of an orthographic character of a variant type font indexed in an electronic document, a pointer to the secondary index corresponding to the character code value of the orthographic character is obtained from the character code value of the orthographic character by referring to the main index table 1 of an index conversion table. The position of entry data to the secondary index of a secondary index table 3 indicated by the obtained pointer is shifted by the variant type number to obtain the font image at the position of the pointer to the head byte of the objective variant type font as an objective variant type font.



LEGAL STATUS

[Date of request for examination] 27.12.1999

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

(19)日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11)特許出願公開番号
特開2000-148736
(P2000-148736A)

(43)公開日 平成12年5月30日 (2000.5.30)

(51)Int.Cl.
G 0 6 F 17/21

識別記号

F I
G 0 6 F 15/20

テマコード(参考)
5 6 2 P 5 B 0 0 9
5 9 6 B

審査請求有 請求項の数12 O L (全11頁)

(21)出願番号 特願平10-321630
(22)出願日 平成10年11月12日 (1998.11.12)

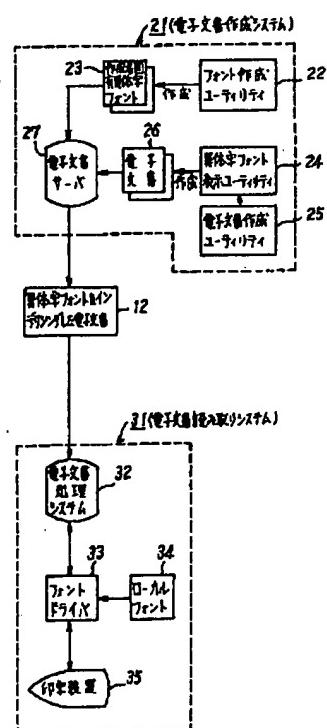
(71)出願人 390009531
インターナショナル・ビジネス・マシーンズ・コーポレーション
INTERNATIONAL BUSINESS MACHINES CORPORATION
アメリカ合衆国10504、ニューヨーク州
アーモンク (番地なし)
(72)発明者 木戸 彰夫
神奈川県大和市下鶴間1623番地14 日本アイ・ピー・エム株式会社 大和事業所内
(74)代理人 100086243
弁理士 坂口 博 (外4名)
Fターム(参考) 5B009 RA00 RA11 RA13 RA14 VC03

(54)【発明の名称】 フォントの取得方法、登録方法、表示方法、印刷方法、異体字フォントを含む電子文書の取り扱い方法およびその記録媒体

(57)【要約】 (修正有)

【課題】 既存の文字符号の文字数の制約を超えて、電子文書の作成者が多くの文字を使用することを可能にするフォントの取得方法、登録方法、表示方法、印刷方法、異体字フォントを含む電子文書の取り扱い方法およびその記録媒体を提供する。

【解決手段】 異体字フォントデータファイルと；主インデックステーブルと、二次インデックステーブルとを備えるインデックス変換テーブルと；からなる異体字フォントファイルを利用する。電子文書中にインデクシングされた対象となる異体字フォントの正字の文字符号値と異体字番号に基づき、正字の文字符号値から正字の文字符号値に対応した二次インデックスへのポインタを得し、取得したポインタで指示された二次インデックステーブルの二次インデックスへのエントリデータの位置を、異体字番号だけ移動して、対象となる異体字フォントの先頭バイトへのポインタ位置のフォントイメージを、対象となる異体字フォントとして取得する。



【特許請求の範囲】

【請求項1】 異体字フォントのイメージデータからなる異体字フォントデータファイルと；正字の文字符号値とその正字の文字符号値に対応した二次インデックスへのポインタとからなる主インデックステーブルと、異体字フォントデータファイル中の対象フォントの先頭バイトへのポインタと、主インデックスが同じ次の二次インデックスエントリへのポインタとから構成される二次インデックスのエントリデータからなる二次インデックステーブルとを備えるインデックス変換テーブルと；からなる異体字フォントファイルを利用し、

電子文書中にインデクシングされた対象となる異体字フォントの正字の文字符号値と異体字番号に基づき、インデックス変換テーブルの主インデックステーブルを参照して、正字の文字符号値から正字の文字符号値に対応した二次インデックスへのポインタを取得し、取得したポインタで指示された二次インデックステーブルの二次インデックスへのエントリデータの位置を、異体字番号だけ移動して、対象となる異体字フォントの先頭バイトへのポインタ位置のフォントイメージを、対象となる異体字フォントとして取得することを特徴とするフォントの取得方法。

【請求項2】 異体字フォントのイメージデータからなる異体字フォントデータファイルと；正字の文字符号値とその正字の文字符号値に対応した二次インデックスへのポインタとからなる主インデックステーブルと、異体字フォントデータファイル中の対象フォントの先頭バイトへのポインタと、主インデックスが同じ次の二次インデックスエントリへのポインタとから構成される二次インデックスのエントリデータからなる二次インデックステーブルとを備えるインデックス変換テーブルと；からなる異体字フォントファイルに異体字フォントを追加するにあたり、

利用者が指定した異体字フォントファイルの異体字フォントデータファイルに追加すべき異体字フォントのイメージを表すフォントデータを、異体字フォントデータファイルにシーケンシャルに追加し登録し、登録後、インデックス変換テーブルにおいて、登録したフォントデータの位置情報、および、フォントファイル中で同じ主インデックスを持つフォントの数を二次インデックスへのエントリデータとして二次インデックステーブルに格納するとともに、利用者が定義した正字の文字符号値とその正字の文字符号値に対応した二次インデックスへのポインタを主インデックステーブルに格納することを特徴とするフォントの登録方法。

【請求項3】 利用者から正字の文字符号値の入力を受けた場合、

フォントライブラリーを検索して、入力された文字符号値を有するフォントを表示装置上に表示するとともに、異体字フォントのイメージデータからなる異体字フォン

トデータファイルと；正字の文字符号値とその正字の文字符号値に対応した二次インデックスへのポインタとからなる主インデックステーブルと、異体字フォントデータファイル中の対象フォントの先頭バイトへのポインタと、主インデックスが同じ次の二次インデックスエントリへのポインタとから構成される二次インデックスのエントリデータからなる二次インデックステーブルとを備えるインデックス変換テーブルと；からなる異体字フォントファイルから、入力された正字の文字符号値を有する異体字フォントを取得して表示装置上に表示することを特徴とするフォントの表示方法。

【請求項4】 表示したフォント群の中から利用者が指定した特定のフォントを、フォントの表示を呼び出したプログラムに返す請求項3記載の異体字フォントを含む電子文書の取り扱い方法。

【請求項5】 フォントを印刷するにあたり、正字については正字の文字符号値を有するフォントをフォントライブラリーから取得して印刷するとともに、異体字フォントのイメージデータからなる異体字フォントデータファイルと；正字の文字符号値とその正字の文字符号値に対応した二次インデックスへのポインタとからなる主インデックステーブルと、異体字フォントデータファイル中の対象フォントの先頭バイトへのポインタと、主インデックスが同じ次の二次インデックスエントリへのポインタとから構成される二次インデックスのエントリデータからなる二次インデックステーブルとを備えるインデックス変換テーブルと；からなる異体字フォントファイルを利用し、

電子文書中にインデクシングされた対象となる異体字フォントの正字の文字符号値と異体字番号に基づき、インデックス変換テーブルの主インデックステーブルを参照して、正字の文字符号値から正字の文字符号値に対応した二次インデックスへのポインタを取得し、取得したポインタで指示された二次インデックステーブルの二次インデックスへのエントリデータの位置を、異体字番号だけ移動して、対象となる異体字フォントの先頭バイトへのポインタ位置のフォントイメージを取得して印刷する、よう構成されたフォントドライバーを使用することを特徴とするフォントの印刷方法。

【請求項6】 請求項1～5のいずれか1項に記載のフォントの取得方法、登録方法、表示方法あるいは印刷方法において、電子文書中のフォントの指定を行う制御シーケンス中に、対象となる異体字フォントがローカルシステム中に存在するのか、オンラインフォントとしてその電子文書中に埋め込まれているのか、それともリモートシステムに存在するのかを指定することを特徴とする電子文書の取り扱い方法。

【請求項7】 電子文書が、HTML、XMLで表示されている場合、電子文書の〈font〉タグにロケーション・キーワードを追加し、ここに存在位置を記憶する請求

項6記載の異体字フォントを含む電子文書の取り扱い方法。

【請求項8】 電子文書中のフォントの指定を行う制御シーケンスを解釈し、指定された異体字フォントがローカルシステムにある場合は、そのローカルシステムから異体字フォントファイルを取得し、そこから異体字フォントのイメージを得る請求項6または7記載の異体字フォントを含む電子文書の取り扱い方法。

【請求項9】 電子文書中のフォントの指定を行う制御シーケンスを解釈し、指定された異体字フォントがインラインフォントである場合は、埋め込まれているインラインフォントの復号を行い、復号したインラインフォントのデータから異体字フォントのイメージを得る請求項6または7記載の異体字フォントを含む電子文書の取り扱い方法。

【請求項10】 電子文書中のフォントの指定を行う制御シーケンスを解釈し、指定された異体字フォントがリモートシステムに存在する場合、フォントリソースを示すURL (Universal Resource Locator) を参照して、そのURLで指定された異体字フォントを取得し、そこから異体字フォントのイメージを得る請求項6または7記載の異体字フォントを含む電子文書の取り扱い方法。

【請求項11】 電子文書中の文字列を検索する際、正字と同じ文字符号値を有する異体字フォントをも検索結果に反映させる請求項6～10のいずれか1項に記載の異体字フォントを含む電子文書の取り扱い方法。

【請求項12】 異体字フォントを、対応する正字の文字符号値と、異体字フォントファイル名と、異体字番号とでインデクシングした電子文書を記録したことを特徴とする記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、複数の情報システム間で情報交換が可能かつ検索可能な方法で異体字を電子文書システムで取り扱うことができる、フォントの取得方法、登録方法、表示方法、印刷方法、異体字フォントを含む電子文書の取り扱い方法およびその記録媒体に関するものである。

【0002】

【従来の技術】 従来から、戸籍管理、顧客管理等の人名・地名を取り扱うアプリケーション、および、古典等の文芸書の電子文書化・出版を行うためのアプリケーションでは、できるだけ多くの文字を取り扱いたいという要求が強かった。その要求に答えるための一つの方法として、ワープロ等の既存のシステムでは、「外字」もしくは「利用者定義文字」という概念を導入し、利用者環境において、その利用者独自の文字を作成・利用して要求に答えていた。

【0003】 しかしながら、この手法は一利用者もしくは一企業等の共通の外字もしくは利用者定義文字を共有

する閉じた環境でのみ利用可能であり、インターネット等のオープンな環境ではそのような外字を使用した文書の情報交換は保証されなかった。つまり、電子文書の参照者の環境では、その文書の作成者が全く意図しなかった文字（字形）が表示されるという結果となった。

【0004】 また、外字を利用した文書の検索には大きな問題が生じていた。たとえば、高島という文字の高にいわゆる梯子高すなわち

【外1】

「高」

を使用したい場合は、その

【外2】

「高」

を外字として作成して使用する。この場合、外字の

【外3】

「高」

20 は正字すなわち標準字形の「高」を用いた検索ではヒットしなかった。その両方の「高」の検索を行おうとする場合、論理和を用いた検索を行うか、または、異体字関係を定義したシソーラスを作成し、そのシソーラスを用いた曖昧検索を行う必要がある。しかしながら、そのどちらの方法も現実的ではなかった。

【0005】 すなわち、論理和を用いた検索を行うためには、利用者はその情報システムに外字として登録されている文字と、それらの文字と標準の文字との異体字関係を全て把握している必要がある。また、シソーラスを用いた手法では、外字は利用者が定義・使用するものであるため、シソーラスに定義されるべき文字は利用者ごとに異なり、共通のシソーラスを検索システムの提供者が作成し利用者に提供することはできなかった。

【0006】 上記の問題を解決するために、必要とされる文字を全て標準の文字として文字符号表上に定義てしまい、さらに、定義された文字の異体関係を定義したシソーラスを作成するという手法が考えられる。しかしながら、字体の細かい違いまでも区別できるように文字符号を定義しようとすると、必要とされる文字数は数十万とも無限とも言われており、その膨大な文字数を持つ文字符号を標準文字符号として全ての情報システム上でサポートすることは、情報システムの利用者全てにその実装コストを転嫁することとなり実質的ではなかった。

また、数十万オーダーの漢字に対して異体字シソーラスを作成した場合、そのサイズは無視できないほどに大きくなり、検索速度に大きな影響を与えるであろうし、文字の追加に際してそのシソーラスを保守することは実質的に無理であった。

【0007】

50 【発明が解決しようとする課題】 上述した問題は、そも

そも漢字という文字に造字機能があることにより発生している。すなわち、手書きの場合、もしくは、活字を用いた通常の出版の場合、必要に応じて字形を新たに作成し、字形の僅かな違いを人間が無視することによって、または、作成した字形にルビを振ることによって、正字と異体字の間の暗黙の関係を作成者と判読者が共有していた。この漢字の造字機能および曖昧な判読機能を情報システム上で実現しようとすると、外字といわれる利用者による文字の定義機能を提供することにより、漢字の造字機能は実現することができる。しかしながら、曖昧な判読機能は、既存の情報システムが曖昧な画像認識処理を行うには非力であったため、および、画像による情報交換では文字符号を用いた情報交換に比べて大きな情報量が必要とされるため、無視されてきた。

【0008】インターネットの普及により、個人・企業の枠を超えて電子的な情報交換が一般的に行われるようになる以前は、印字された文字を媒体として情報交換を行っていたため、判読は利用者の責任においてなされ、情報システム側に曖昧な判読機能を持たせようとする要求が低かった。しかし、電子的な情報交換、および、電子的に蓄積された情報の検索が一般的に行われるようになってきた昨今、標準として規定されている情報交換可能な文字符数の不足の問題、外字として造字された文字の情報交換の必要性、および、検索への要求が高くなっている。

【0009】本発明の目的は上述した課題を解消して、既存の文字符号の文字符数の制約を超えて、電子文書の作成者が多くの文字を使用することを可能にする、フォントの取得方法、登録方法、表示方法、異体字フォントを含む電子文書の取り扱い方法およびその記録媒体を提供しようとするものである。

【0010】

【課題を解決するための手段】本発明のフォントの取得方法では、異体字フォントのイメージデータからなる異体字フォントデータファイルと；正字の文字符号値とその正字の文字符号値に対応した二次インデックスへのポインタとからなる主インデックステーブルと、異体字フォントデータファイル中の対象フォントの先頭バイトへのポインタと、主インデックスが同じ次の二次インデックスエントリへのポインタとから構成される二次インデックスのエントリデータからなる二次インデックステーブルとを備えるインデックス変換テーブルと；からなる異体字フォントファイルを利用する。

【0011】異体字フォントファイルの利用方法は以下の通りである。電子文書中にインデクシングされた対象となる異体字フォントの正字の文字符号値と異体字番号に基づき、インデックス変換テーブルの主インデックステーブルを参照して、正字の文字符号値から正字の文字符号値に対応した二次インデックスへのポインタを取得し、取得したポインタで指示された二次インデックステ

ーブルの二次インデックスへのエントリデータの位置を、異体字番号だけ移動して、対象となる異体字フォントの先頭バイトへのポインタの位置のフォントイメージを、対象となる異体字フォントとして取得する。

【0012】また、本発明のフォントの登録方法では、異体字フォントファイルに異体字フォントを追加するにあたり、利用者が指定した異体字フォントファイルの異体字フォントデータファイルに追加すべき異体字フォントのイメージを表すフォントデータを、異体字フォントデータファイルにシーケンシャルに追加し登録し、登録後、インデックス変換テーブルにおいて、登録したフォントデータの位置情報、および、フォントファイル中で同じ主インデックスを持つフォントの数を二次インデックスへのエントリデータとして二次インデックステーブルに格納するとともに、利用者が定義した正字の文字符号値とその正字の文字符号値に対応した二次インデックスへのポインタを主インデックステーブルに格納する。

【0013】さらに、本発明のフォントの表示方法では、利用者から正字の文字符号値の入力を受けた場合、まず、フォントライブラリーを検索して、入力された文字符号値を有するフォントを表示装置上に表示する。同時に、異体字フォントのイメージデータからなる異体字フォントデータファイルと；正字の文字符号値とその正字の文字符号値に対応した二次インデックスへのポインタとからなる主インデックステーブルと、異体字フォントデータファイル中の対象フォントの先頭バイトへのポインタと、主インデックスが同じ次の二次インデックスエントリへのポインタとから構成される二次インデックスのエントリデータからなる二次インデックステーブルとを備えるインデックス変換テーブルと；からなる異体字フォントファイルから、入力された正字の文字符号値を有する異体字フォントを取得して表示装置上に表示する。

【0014】さらにまた、本発明のフォントの印刷方法では、まず、フォントを印刷するにあたり、正字については正字の文字符号値を有するフォントをフォントライブラリーから取得して印刷する。同時に、異体字フォントのイメージデータからなる異体字フォントデータファイルと；正字の文字符号値とその正字の文字符号値に対応した二次インデックスへのポインタとからなる主インデックステーブルと、異体字フォントデータファイル中の対象フォントの先頭バイトへのポインタと、主インデックスが同じ次の二次インデックスエントリへのポインタとから構成される二次インデックスのエントリデータからなる二次インデックステーブルとを備えるインデックス変換テーブルと；からなる異体字フォントファイルを利用し、電子文書中にインデクシングされた対象となる異体字フォントの正字の文字符号値と異体字番号に基づき、インデックス変換テーブルの主インデックステーブルを参照して、正字の文字符号値から正字の文字符号

値に対応した二次インデックスへのポインタを取得し、取得したポインタで指示された二次インデックステーブルの二次インデックスへのエントリデータの位置を、異体字番号だけ移動して、対象となる異体字フォントの先頭バイトへのポインタ位置のフォントイメージを取得して印刷する。

【0015】また、上述したフォントの取得方法、登録方法、表示方法、印刷方法において、異体字フォントを含む電子文書の取り扱い方法の好ましい態様として、異体字フォントを電子文書中に埋め込む方法と、異体字フォントを埋め込んだ電子文書の処理方法と、異体字フォントを埋め込んだ電子文書中の文字列の検索方法とを提供する。さらにまた、本発明例およびその好適例で使用する異体字フォントをインデクシングした電子文書を記憶した記録媒体を提供する。

【0016】上述したフォントの取得方法、登録方法、表示方法、印刷方法、異体字フォントを含む電子文書の取り扱い方法では、既存の文字符号の文字数の制約を超えて、電子文書の作成者が多くの文字（字形）を使用することを可能にする。また、そのことによって、戸籍管理、顧客管理等の人名・地名を取り扱うアプリケーション、および、古典等の文芸書の電子文書化・出版等のアプリケーションが使用できる文字数の制約を取り除き、正確な字形を使用した人名、地名、文芸書の処理および蓄積を可能にする。さらに、利用者によって造字された文字の電子文書による情報交換を可能にする。すなわち、電子文書の受信者の環境において、その作成者の環境と同じ字形で造字された文字を表示可能にする。さらにまた、造字された文字を用いた電子文書の検索を可能とする。すなわち、正字と異体字を区別せずに文字列の検索を行う、もしくは、必要に応じて正字と異体字を区別さずに文字列の検索を行うことができる。

【0017】また、本発明の好ましい態様として、電子文書中への異体字フォントの埋め込みを、ローカルシステムにおける文字符号、インラインフォントとして符号化された文字符号、または、リモートシステム上のフォントファイルの文字符号を示すURLを使用することを実施する。この場合は、既存の文字符号を用いて上記の要求を満たす電子文書処理システムを実現することにより、実装コストを削減することができる。

【0018】

【発明の実施の形態】まず、本発明の異体字フォントの電子文書中へのインデクシングについて説明する。本発明では、異体字フォントを、対応する正字の文字符号値と、異体字フォントファイル名と、異体字番号とで、電子文書中へ指定する。そのため、インデクシングされた異体字フォントは、通常のフォントがフォント名および文字符号による一次元のインデックスを持つのに対し、二次元のインデックスから構成されている。この二次元のインデックスは、後述するようにフォントファイルの

中に内蔵されたインデックス変換テーブルを参照することによって、フォントファイル中のフォントの位置情報を翻訳される。ここで、副インデックスの異体字番号は、好ましくは、対象フォントのフォントファイルにおいて、フォントデータファイルの先頭からのフォント数からなるシーケンシャルなフォント番号、または、対象フォントデータの先頭バイトの位置を示すバイトインデックスである。

【0019】異体字フォントの異体字番号（副インデックス）は、フォント名の中に符号化されるものとする。

10 フォント名は、タイプフェースを示す情報、作成者を示す情報、異体字番号からなる。例えば、“MINCHO-ITAI@FOOBAR:2”という構造を持つ。この場合、“MINCHO-ITA”はタイプフェース名、“@”はタイプフェースと作成者情報を区切る区切り文字、“FOOBAR”は作成者情報、“：“は作成者情報と異体字番号を区切る区切り文字、“2”は異体字番号である。異体字フォントは、このフォント名とこのフォント名の後に続く正字の文字符号とから構成される。

20 【0020】次に、インデックス変換テーブルを参照して上述した構成の異体字フォントをインデクシングする方法の一例を説明する。図1は異体字フォントのインデックス変換テーブルの一例の構成を示す図、図2は図1に示すインデックス変換テーブルを利用して異体字フォントをインデクシングする方法の一例を説明するためのフローチャートである。図1において、インデックス変換テーブルは、主インデックステーブル1と二次インデックステーブル3とから構成されている。主インデックステーブル1は、正字の文字符号値とその正字符串号値に

30 対応した二次インデックスへのポインタとから構成される。二次インデックステーブル3は、異体字フォントデータファイル中の対象フォントの先頭バイトへのポインタと、主インデックスが同じ次の二次インデックスエントリへのポインタとから構成される二次インデックスへのエントリデータ2から構成される。なお、4は、フォントファイル中の異体字フォントからなるフォントデータファイルである。このインデックス変換テーブルは、各利用者のフォントファイル中に内蔵されてフォントドライバが認識できるよう構成する。

40 【0021】図2に従って、図1に示すインデックス変換テーブルを利用して異体字フォントをインデクシングする方法を説明する。まず、電子文書中にインデックスされた異体字フォントから正字の文字符号値を読み取り、主インデックステーブル1で利用する正字符串号値番号のエントリを得る（S1）。主インデックステーブル1を参照することで、得られたエントリの内容から二次インデックステーブル3へのポインタを求める（S2）。この二次インデックステーブル3へのポインタは、異体字フォントの主インデックスで定義された正字のうち0番目のエントリとなる。次に、異体字フォント

の副インデックスから得られた異体字番号で表される目的の二次インデックス値と、エントリの番目が一致するかどうかを検討する(S3)。目的の二次インデックス値とエントリの番目が一致した場合は、二次インデックスステーブル3中の二次インデックスへのエントリデータ2から、フォントデータファイル4における対象フォントの先頭バイトへのポインタを得る(S4)。そして、得られた先頭バイトのポインタからフォントデータファイル4中の対象となるフォントをアクセスして表示等をする(S5)。

【0022】一方、目的の二次インデックス値とエントリの番目とが一致しない場合は、二次インデックスへのエントリデータ2の中の主インデックスが同じ次の二次インデックスエントリへのポインタを読み取り、読み取ったポインタがナル(null)かどうかを検討する(S6)。読み取ったポインタがナルの場合は、次のエントリが見つからないため、エラーを表示して(S7)終了する。読み取ったポインタがナルでない場合は、得られている二次インデックスへのエントリへのポインタを得、得られたポインタから主インデックスが同じ次の番目のエントリを得る(S8)。以後、ステップS3以降のステップを繰り返すことで、異体字フォントのインデクシングが終了する。なお、上述した例では、二次インデックスへのエントリデータ2として、フォントデータファイル4に対する対象フォントの先頭バイトへのポインタを利用したが、それに代えて、フォントデータファイル4に対する先頭からのフォント数からなるシーケンシャルなフォント番号を利用することができる。

【0023】次に、本発明の異体字フォントを含む電子文書の取り扱い方法を実施するシステムの一例について、図3を参照して説明する。図3に示す例において、電子文書処理システム11は、電子文書を作成するための電子文書作成システム21と、電子文書を読み取るための電子文書読み取りシステム31とから構成されている。図3に示す例では、電子文書読み取りシステム31からの要求で、電子文書作成システム21で作成した電子文書を閲覧しようとする際、上述した構成の本発明の異体字フォントをインデクシングした電子文書12を利

用する。

【0024】電子文書作成システム11には、異体字フォント12を作成するためのフォント作成ユーティリティ22と作成者固有の異体字フォントファイル23、および、異体字を含む電子文書を作成するために使用するフォント表示ユーティリティ24とワードプロセッサ等の電子文章作成ユーティリティ25と作成した電子文書ファイル26、さらに、異体字フォントファイル23と電子文書ファイル26を蓄積する電子文書サーバ27を備えている。また、電子文書読み取りシステム31には、異体字フォントを含む電子文書を読み取るための電子文書処理システム32、および、フォントドライバ3

3、ローカルフォント34、印刷装置35を備えている。

【0025】以下、図3に示す電子文書処理システム1のうち、異体字フォントの作成ユーティリティ23、異体字フォントの表示ユーティリティ24、および、異体字フォントを取り扱うフォントドライバ33の構成について説明する。

【0026】異体字フォントの作成ユーティリティ23は、利用者によって入力された、タイプフェース名と作成者の名前を持つフォントファイル中に(対応するフォントファイルがなければ作成する)、上述した本発明の構成を有する異体字フォントを追加する。異体字フォントの字形データは、フォントファイルのフォントデータファイル4にシーケンシャルに追加されるものとする。字形データの登録後、異体字フォントの作成ユーティリティ23はインデックス変換テーブルに新しいエントリを追加する。そのエントリに、登録した字形データの位置情報と、利用者によって入力された対応する正字の文字符号(主インデックス)、および、上記フォントファイル中で同じ主インデックスを持つフォントの数を副インデックスとして格納する。これらの操作は、上述したインデックス変換テーブルを利用した異体字フォントのインデックス方法を参照することができる。

【0027】なお、異体字フォントの主インデックスの指定は、通常の外字作成ユーティリティの場合と同じく、フォントデータの作成者に委ねられる。このインデックスの指定により、作成された異体字は正字と対応づけられ、検索の際に利用されることとなる。この対応関係は文字符号として電子文書中に含まれるので、検索の際に曖昧検索のためのシソーラスを参照する必要がなくなる。また、正字と異体字の関係は、異体字の作成者が定義したものである。そのため、その電子文書の検索において、作成者側と閲覧者側の間で解釈の違いによる検索結果の違いが生じることもなくなる。

【0028】異体字フォントの表示ユーティリティ24は、利用者から正字のフォント名および文字符号の入力を受け、情報システムによりサポートされるフォントライブラリを検索して、上記フォント名および文字符号を主インデックスとして持つフォントを全て表示装置上に表示する。さらに、キーボードまたはポインティングデバイスを通じて利用者からの表示したフォント群の中から特定のフォントを示す入力を得て、対応するフォントのフォント名を利用者もしくは異体字フォントの表示ユーティリティ24を呼び出したプログラムに返すものとする。本発明の好適例においては、異体字の表示ユーティリティ24は、かな漢字変換等のInput Method Editor(IME)によって呼び出される。IMEは、さらにワードプロセッサ等のIMEを呼び出したアプリケーションプログラムに、主インデックス(正字のフォント名および文字符号)と副インデックス(異体字番号)からなる本發

11

明の異体字フォントを返すことにより、アプリケーションは電子文書中に指定された異体字フォントのフォント情報と正字の文字符号とを埋め込むことができる。

【0029】図4に示すフローチャートに従って異体字フォントを取り扱うフォントドライバ33を説明する。まず、フォントドライバ33は、指定されたフォント名中に異体字番号を区切る区切り文字が存在するかどうかを検討する(S11)。区切り文字が存在しない場合は、通常処理に戻る(S12)。区切り文字が存在する場合は、先頭からその区切り文字の直前の文字までをフォント名として、システムがサポートするフォントライブラリを検索し、該当するフォントファイルを得る(S13)。さらに、上記区切り文字の直後の文字から最後の文字までの文字列を数値に変換し、それを異体字番号とする。そして、その異体字番号を副インデックスとし(S14)、指定された正字のフォント名およびそのフォント名に続く文字符号を主インデックスとする(S15)。次に、異体字フォントファイル中のインデックス変換テーブルを参照し、正・副からなる二次元のインデックスをフォントの位置情報に変換し、アクセスすべきフォント(字形データ)を得(S16)、表示装置上にそのフォントを印字する(S17)。このインデックス変換テーブルを参照する方法はすでに説明した方法と同じである。

【0030】次に、本発明の好適例として、異体字フォントを電子文書中に埋め込む方法および異体字フォントを埋め込んだ電子文書の所定方法について説明する。本発明では、これから説明する異体字フォントの情報は、上述した本発明の異体字フォントを含む電子文書の取り扱い方法に付加して利用されている。まず、図3に示す例において、電子文書作成ユーティリティ25では、電子文書中のフォントの指定を行う制御シーケンスにおいて、そのフォントのフォント名および文字符号とともに、そのフォントが、ローカルシステム中に存在するのか、オンラインフォントとしてその電子文書中に埋め込まれているのか、あるいは、リモートシステムに存在するのかを示すフォントロケーションを指定する。このフォントロケーションの指定は、例えば、HTML、XML等での実装の場合、電子文書中の〈font〉タグにlocationキーワードを追加し、そのlocationキーワードは値として、ローカルシステムを示すLOCAL、オンラインフォントを示すINLINE、リモートシステム上のフォントファイルを示すURL(Universal Resource Locator)が持てるようになることで行う。

【0031】異体字フォントを埋め込んだ電子文書の処理システム32では、処理を行う電子文書のフォーマット中のフォントの指定を行う制御シーケンスのフォントロケーションを解釈し、指定されたフォントが、ローカルシステムに存在するのか(通常の正字フォントの場合)、その電子文書に埋め込まれたオンラインフォント

12

であるのか、それともリモートシステムに存在するのかを知る。

【0032】参照されているフォントがローカルシステムに存在する場合は、各利用者のシステムに存在する電子文書作成者と同じ通常のフォントを参照し、通常の処理で文字の印字を行う。

【0033】参照されているフォントがオンラインフォントであった場合、電子文書の処理システム32は、埋め込まれているオンラインフォントの復号を行い、電子文書の処理システム32が獲得した一時ストレージの中に異体字フォントのバイナリイメージを展開する。そして、展開されたフォントファイル、電子文書中で参照されているフォント名、文字符号(その異体字に対応する正字の文字符号)をフォントドライバ33に引き渡すことにより、その異体字の印字を行う。ここで、MIME(Multi-purpose Internet Mail Extension)のメディアタイプとして定義されたオンラインフォントには、インターネットで交換可能な電子文書の中にフォントが含まれるようにする。サブタイプとしては、TrueType、Type 1等の既存のフォントフォーマットおよび本発明が提案する複数のフォントファイルからなる複合フォントを示すフォーマットを指定できるようにする。オンラインフォントとして電子文書中に埋め込まれるフォントは、好ましくはBase64符号化法によって符号化される。

【0034】参照されているフォントがリモートシステムに存在する場合、フォントリソースを示すURLをインターネットを介して参照し、電子文書の処理システムが獲得した一時ストレージの中に異体字フォントのバイナリイメージをコピーする。そして、コピーされたフォントファイル、電子文書中で参照されているフォント名、文字符号をオンラインフォントの場合と同じようにフォントドライバ33に引き渡すことにより、その文字の印字を行う。

【0035】なお、上述した本発明の好適例では、異体字フォントを所定の構成にインデクシングした電子文書を対象として説明したが、オンラインフォントを利用する場合は上述した本発明の異体字フォントのインデクシングによらなくでも異体字フォントを取り扱えることはいうまでもない。

【0036】本発明では、電子文書中において、異体字は対応する正字の文字符号をもって表現される。しかしながら、その電子文書の印字に際しては、その電子文書の作成者によって作成された異体字フォントが参照されるため、印字字形としては作成者が意図した異体字がどの環境でも表示されることとなり、異なるシステム間ににおける情報交換によって文字化けが起こることがなくなり、印字字形についても正確な情報交換が可能となる。

【0037】最後に、異体字フォントを埋め込んだ電子文書中の文字列の検索を行う手法について説明する。本発明の異体字フォントを含む電子文書の取り扱い方法で

は、電子文書中にインデクシングされた異体字フォントは文字符号としては対応する正字と同じ符号値をもっている。そのため、通常のフォント情報を無視した検索の場合、正字に加えて、その正字に結びつけられた異体字もヒットすることになる。多くの場合、正字の検索の結果として正字と異体字の両方がヒットすることは、簡単な検索式により利用者が意図した文字列がヒットすることになり、検索として都合がよい。利用者が正字と異体字とを区別した検索を行いたい場合は、利用者は文字符号に加えて、フォント名を指示することで、検索エンジンに電子文書中の文字列の検索を行わせることができると。

【0038】

【発明の効果】以上の説明から明らかのように、本発明のフォントの取得方法、登録方法、表示方法、印刷方法、異体字フォントを含む電子文書の取り扱い方法およびその記録媒体では、異体字フォントをインデクシングした電子文書を所定の方法で取り扱うことで、既存の文字符号の文字数の制約を超えて、電子文書の作成者が多くの文字（字形）を使用することを可能にする。また、そのことによって、戸籍管理、顧客管理等の人名・地名を取り扱うアプリケーション、および、古典等の文芸書の電子文書化・出版等のアプリケーションが使用できる文字数の制約を取り除き、正確な字形を使用した人名、地名、文芸書の処理および蓄積を可能にする。さらに、利用者によって造字された文字の電子文書による情報交換を可能にする。すなわち、電子文書の受信者の環境において、その作成者の環境と同じ字形で造字された文字を表示可能にする。さらにまた、造字された文字を用いた電子文書の検索を可能とする。すなわち、正字と異体字

30

を区別せずに文字列の検索を行う、もしくは、必要に応じて正字と異体字を区別せずに文字列の検索を行うことができる。

【0039】また、本発明の異体字フォントを含む電子文書の取り扱い方法の好適例では、既存の文字符号を用いて上記の要求を満たす電子文書処理システムを実現することにより、実装コストを削減することができる。

【図面の簡単な説明】

【図1】本発明の異体字フォントのインデックス変換テーブルの一例の構成を示す図である。

【図2】図1に示すインデックス変換テーブルを利用して異体字フォントをインデクシングする方法を説明するためのフローチャートである。

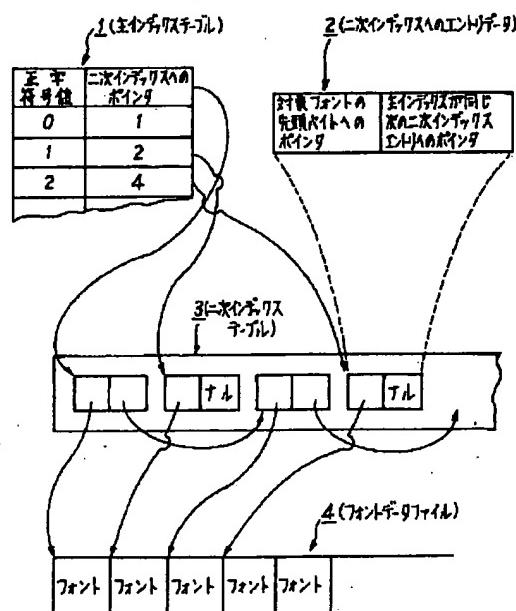
【図3】本発明の電子文書処理システムにおける異体字フォントの取り扱い方法の第1発明を説明するための図である。

【図4】異体字フォントを取り扱うフォントドライバにおける異体字フォントの取り扱い方法を説明するためのフローチャートである。

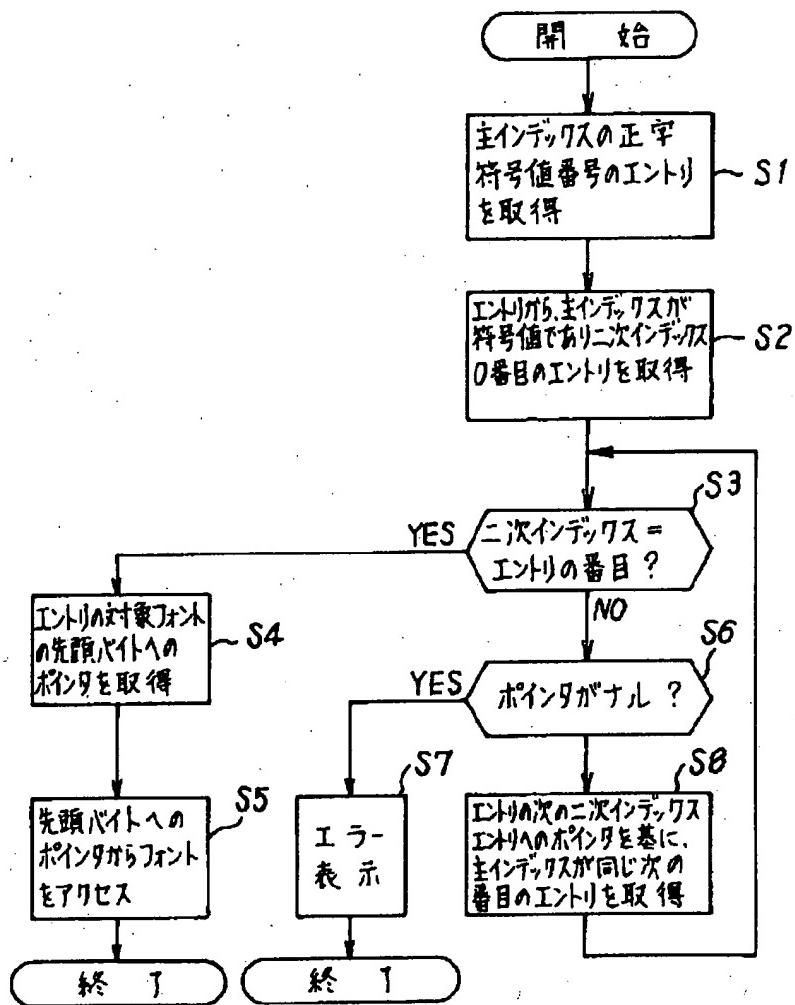
【符号の説明】

- 1 主インデックステーブル、2 二次インデックスへのエントリデータ、3 二次インデックステーブル、4 フォントデータファイル、11 電子文書処理システム、12 異体字フォント、13 異体字フォントデータ、21 電子文書作成システム、22 フォント作成ユーティリティ、23 異体字フォントファイル、24 フォント表示ユーティリティ、25 電子文書作成ユーティリティ、26 電子文書ファイル、27 電子文書サーバ

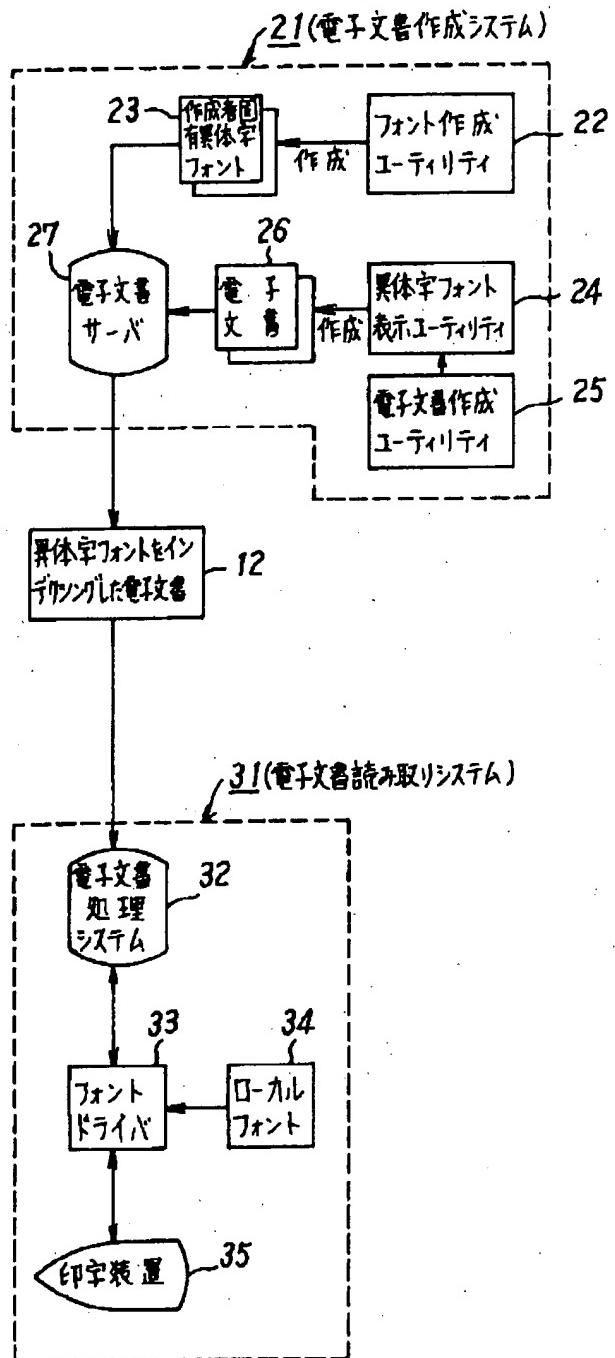
【図1】



【図2】



【図3】



【図4】

